Fairness with an Honest Minority and a Rational Majority

Fairness with an Honest Minority and a Rational Majority*

Shien Jin Ong† David C. Parkes‡ Alon Rosen§ Salil Vadhan¶

December 21, 2008

Abstract

We provide a simple protocol for secret reconstruction in any threshold secret sharing scheme, and prove that it is *fair* when executed with many *rational* parties together with a small minority of *honest* parties. That is, all parties will learn the secret with high probability when the honest parties follow the protocol and the rational parties act in their own self-interest (as captured by a set-Nash analogue of trembling hand perfect equilibrium). The protocol only requires a *standard* (synchronous) broadcast channel, tolerates both early stopping and incorrectly computed messages, and only requires 2 rounds of communication.

Previous protocols for this problem in the cryptographic or economic models have either required an honest majority, used strong communication channels that enable simultaneous exchange of information, or settled for approximate notions of security/equilibria. They all also required a nonconstant number of rounds of communication.

Keywords: game theory, fairness, secret sharing.

*Earlier versions of this paper are [38, 39].

†Goldman, Sachs & Co., New York, NY. E-Mail: shienjin@alum.mit.edu. Work done while author was a graduate student at Harvard School of Engineering and Applied Sciences.

‡Harvard School of Engineering and Applied Sciences, Cambridge, MA. E-Mail: parkes@eecs.harvard.edu

§Herzliya IDC, Israel. E-Mail: alon.rosen@idc.ac.il. Part of this work done while the author was a postdoctoral fellow at Harvard University's Center for Research of Computation and Society (CRCS). Work supported in part by Israel Science Foundation (ISF) Grant no. 334/08.

¶Harvard School of Engineering and Applied Sciences and Center for Research on Computation and Society, Cambridge, MA. E-Mail: salil@eecs.harvard.edu. Supported by NSF grant NSF grant CNS-0831289.

1 Introduction

A major concern in the design of distributed protocols is the possibility that parties may deviate from the protocol. Historically, there have been two main paradigms for modeling this possibility. One is the cryptographic paradigm, where some parties are honest, meaning they will always follow the specified protocol, and others are malicious, meaning they can deviate arbitrarily from the protocol. The other is the economic paradigm, where all parties are considered to be rational, meaning that they will deviate from the protocol if and only if it is in their interest to do so.

Recently, some researchers have proposed studying mixtures of these traditional cryptographic and economic models, with various combinations of honest, malicious, and rational participants. One motivation for this that it may allow a more accurate modeling of the diversity of participants in real-life executions of protocols. Along these lines, the papers of Aiyer et al. [3], Lysyanskaya and Triandopoulos [34], and Abraham et al. [2] construct protocols that achieve the best of both worlds. Specifically, they take protocol properties that are known to be achievable in both the cryptographic model (with honest and malicious parties) and the economic model (with only rational parties), and show that protocols with the same properties can still be achieved in a more general model consisting of malicious and rational parties.

Our work is of the opposite flavor. We consider properties that are not achievable in either the cryptographic or economic models alone, and show that they can be achieved in a model consisting of both honest and rational parties. Specifically, we consider the task of secret reconstruction in *secret sharing*, and provide a protocol that is *fair*, meaning that all parties will receive the output, given rational participants together with a small minority of honest participants. In standard communication models, fairness is impossible in a purely economic model (with only rational participants) [23, 28] or in a purely cryptographic model (with a majority of malicious participants) [13]. Previous works in the individual models achieved fairness by assuming strong communication primitives that allow simultaneous exchange of information [23, 22, 2, 28, 32, 33, 24][1] or settled for approximate notions of security/equilibria [15, 9, 20, 43, 28], whereas we only use a standard (i.e. synchronous but not simultaneous) broadcast channel and achieve a standard notion of game-theoretic equilibrium (namely, a trembling hand perfect equilibrium).

Thus, our work illustrates the potential power of a small number of honest parties to maintain equilibria in protocols. These parties follow the specified strategy even when it is not in their interest to do so, whether out of altruism or laziness. While we study a very specific problem (secret sharing reconstruction, as opposed to general secure function evaluation), we hope that eventually the understanding developed in this clean setting will be leveraged to handle more complex settings (as has been the case in the past).

Below, we review the cryptographic and economic paradigms in more detail. We then introduce the secret-sharing problem we study and survey recent works on this problem in the purely economic model. We then describe our results and compare them to what was achieved before.

1.1 The Cryptographic Paradigm

In the cryptographic paradigm, we allow for a subset of the parties to deviate from the protocol in an arbitrary, malicious manner (possibly restricted to computationally feasible strategies), and the actions of these parties are viewed as being controlled by a single adversary. Intuitively, this

[1] Actually, the impossibility results of [23, 28] also hold in the presence of a simultaneous broadcast channel and thus the works of [23, 22, 2, 28] use additional relaxations, such as allowing the number of rounds and/or the sizes of the shares to be unbounded random variables.

captures worst-case deviations from the protocol, so protocols protecting against such malicious and monolithic adversaries provide a very high level of security. Remarkably, this kind of security can be achieved for essentially every multiparty functionality, as shown by a series of beautiful results from the 1980's [50, 21, 11, 7, 44]. However, considering arbitrary (and coordinated) malicious behavior does have some important limitations. For example, it is necessary to either assume that a majority of the participants are honest (i.e. not controlled by the adversary) or allow for protocols that are unfair (i.e. the adversary can prevent some parties from getting the output). This follows from a classic result of Cleve [13], who first showed that there is no fair 2-party protocol for coin-tossing (even with computational security), and then deduced the general version by viewing a multiparty protocol an interaction between two super-parties, each of which controls half of the original parties. Lepinski et al. [32] bypass this impossibility result by assuming a strong communication primitive ("ideal envelopes") which allow simultaneous exchange of information, but it remains of interest to find ways of achieving fairness without changing the communication model.

1.2 The Economic Paradigm

In the economic paradigm, parties are modeled as rational agents with individual preferences, and will only deviate from the protocol if this is in their own self interest. This approach has become very popular in the computer science literature in recent years, with many beautiful results. There are two aspects of this approach:

1. Design computationally efficient mechanisms (i.e. functionalities that can be implemented by a trusted mediator) that give parties an incentive to be truthful about their private inputs, while optimizing some *social choice function*, which measures the benefit to society and/or the mechanism designer [36, 31, 5].

2. Implement these mechanisms by distributed protocols, with computational efficiency emphasized in *distributed algorithmic mechanism design* [16, 17, 18] and extended to also emphasize additional equilibrium considerations in *distributed implementation* [47, 41, 42], so that parties are "faithful" and choose to perform message passing and computational tasks in *ex post* Nash equilibrium. More recent works achieve a strong form of distributed implementation, with provably no additional equilibria [33, 24], but require strong communication primitives.

Note that distributed algorithmic mechanism design is different in spirit from the traditional problem considered in cryptographic protocols, in that parties have "true" private inputs (whereas in cryptography all inputs are considered equally valid) and there is freedom to change how these inputs are mapped to outcomes through choosing appropriate social choice functions to implement (whereas in cryptography, the functionality is pre-specified.) Nevertheless, recent works have explored whether we can use the economic model to obtain 'better' solutions to traditionally cryptographic problems, namely to compute some pre-specified functionalities. One potential benefit is that we may be able to incentivize parties to provide their "true" private inputs along the lines of Item 1 above; the papers [35, 48] explore for what functionalities and kinds of utility functions this is possible in the presence of game-theoretic agents.

A second potential benefit is that rational deviations may be easier to handle than malicious deviations (thus possibly leading to protocols with better properties), while also preferable to assuming a mixture of players at the honest and malicious extremes. This has led to a line of work, started by Halpern and Teague [23] and followed by [22, 2, 28], studying the problems of secret sharing and multiparty computation in the purely economic model, with all rational participants. One can also require notions of equilibria that are robust against coalitions of rational

players [2]. While this approach has proved to be quite fruitful, it too has limitations. Specifically, as pointed out in [22, 28], it seems difficult to construct rational protocols that are fair in the standard communication model, because parties may have an incentive to stop participating once they receive their own output. The works [23, 22, 2, 28], as well as [33, 24] applied to appropriately designed mediated games, achieve fairness by using strong communication primitives (simultaneous broadcast, "ballot boxes") that allow simultaneous exchange of information.

As mentioned above, we achieve fairness in the standard communication model by considering a mix of rational participants together with a *small* minority of honest participants. Note that Cleve's [13] proof that an honest majority is necessary in the cryptographic setting, by reduction to the two-party case, no longer applies. The reason is that we cannot view a subset of the rational parties as being controlled by a single super-party. Even in coalitional notions of equilibria, each individual in that subset would only agree to a coordinated (joint) deviation if it is in its own interest to do so.

Our protocol is for the share reconstruction problem in secret sharing, which we now describe in more detail.

1.3 Secret Sharing

In a *t-out-of-n secret-sharing scheme* [46, 8], a dealer takes a secret s and computes n (randomized) *shares* $s_1, \ldots, s_n$ of s, which are distributed among n parties. The required properties are that (a) any set of t parties can reconstruct the secret s from their shares, but (b) any set of fewer than t parties has no information about s (i.e. they would have been equally likely to receive the same shares for every possible value of s).

Secret sharing is a fundamental building block for cryptographic protocols [21, 7, 11, 44]. Typically, these protocols are structured as follows. First, every party shares its private input among all the parties. Then the computation of the functionality is done on shares (to maintain privacy). And at the end, the parties reveal their shares of the output so that everyone can reconstruct it. Our focus in this paper is on this final reconstruction step. Typically, it is assumed that there are enough honest parties in the protocol to ensure that the secret can be reconstructed from the revealed shares, even if some parties refuse to reveal their shares or even reveal incorrect values. This turns out to be achievable if and only if more than a 2/3 fraction of the players are honest [12]. (In previous versions of this paper, we restricted attention to *fail-stop deviations* where a party may stop participating in the protocol early but otherwise follows the prescribed strategy, in which case only an honest majority is needed in the traditional cryptographic model.)

1.4 Rational Secret Sharing

It is natural to ask whether we can bypass this need for an honest majority by considering only *rational* deviations from the protocol. As noted above, the study of secret sharing with only rational participants was initiated by Halpern and Teague [23], and there have been several subsequent works [22, 28, 2]. In these works, it is assumed that participants prefer to learn the secret over not learning the secret, and secondarily, prefer that as few other agents as possible learn it. As pointed out in Gordon and Katz [22], any protocol where rational participants reveal their shares sequentially will not yield a Nash equilibrium. This is because it is rational for the t'th player to stop participating, as she can already compute the secret from the shares of the first $t-1$ players and her own, and stopping may prevent the first $t-1$ players from learning it.

One way to get around this difficulty is to assume a *simultaneous broadcast channel*, where all parties can broadcast values at the same time, without the option of waiting to see what values

the other parties are broadcasting. All parties simultaneously revealing their shares is a Nash equilibrium. That is, assuming all of the other parties are simultaneously revealing their shares, no party can increase her utility by aborting (stopping early) instead of revealing. This basic protocol is instructive because it has several deficiences:

1. A simultaneous broadcast channel is a strong (and perhaps unrealistic) communication primitive, particularly in the context of trying to achieve fairness, where the typical difficulties are due to asymmetries in the times that parties get information. For example, fair coin-tossing is trivial with a simultaneous broadcast channel (everyone broadcasts a bit, and the output is the exclusive-or), in contrast to Cleve's impossibility result for synchronous broadcast channels [13].

2. Nash Equilibrium in this context is a very weak guarantee. For example, as argued by Halpern and Teague [23], it seems likely that rational parties would actually abort. The reason is that aborting is never worse than revealing, and is sometimes better (if $t-1$ other parties reveal, then the tth party will always learn the secret and can prevent the other parties from doing so by an abort.)

Halpern and Teague [23] and follow-up works [22, 2, 28] focus on the second issue. That is, they allow simultaneous broadcast, and explore whether stronger solution concepts than plain Nash equilibrium can be achieved. Halpern and Teague [23] propose looking for an equilibrium that survives "iterated deletion of weakly dominated strategies." They prove that no bounded-round protocol can achieve a fair outcome in equilibrium when adopting this solution concept. However, they and subsequent works by Gordon and Katz [22] and Abraham et al. [2] show that fair outcomes are possible even with this equilibrium refinement using a probabilistic protocol whose number of rounds has finite expectation. Moreover, Abraham et al. [2] show how to achieve an equilibrium that is resistant to deviations by coalitions of limited size. Kol and Naor [28] argue that "strict equilibria" is a preferable solution concept to the iterated deletion notion used by Halpern and Teague [23], and show how to achieve it with a protocol where the size of shares dealt is an unbounded random variable with finite expectation. (They also show that a strict equilibrium cannot be achieved if the shares are of bounded size.) In all of the above works, the protocols' prescribed instructions crucially depend on the utilities of the various players.

The works of Lepinski et al. [33] and Izmalkov et al. [24] also can be used to obtain fair protocols for secret sharing by making an even stronger physical assumption than a simultaneous broadcast channel, namely "ballot boxes." Specifically, they show how to compile any game with a trusted mediator into a fair ballot-box protocol with the same incentive structure. Since the share-reconstruction problem has a simple fair solution with a trusted mediator (the mediator takes all the inputs, and broadcasts the secret iff *all* players reveal their share), we can apply their compiler to obtain a fair ballot-box protocol. But our interest in this paper is on retaining standard communication models.

1.5 Our Results

In this paper, we address both issues above. Specifically, we assume that there is a *small* number k of honest participants (which can be much smaller than the secret-sharing threshold t), and the rest are rational. In this setting, we exhibit a simple protocol that only requires a standard communication model, namely *synchronous broadcast*, and in cases where the total number of players is sufficiently large, achieves fair outcomes with high probability with respect to a strong

solution concept, namely (a set-Nash analogue of) *trembling hand perfect equilibrium*. We describe both aspects of our result in more detail below.

Synchronous Broadcast. With a *synchronous* (as opposed to *simultaneous*) broadcast channel, the protocol proceeds in rounds, and only one party can broadcast in each round.[2] When all parties are rational, the only previous positive results in this model are in works by Kol and Naor [28, 27], who achieve a fair solution with an approximate notion of Nash equilibrium — no party can improve her utility by ε by deviating from the protocol. However, it is unclear whether such ε-Nash equilibria are satisfactory solution concepts because they may be unstable. In particular, how can everyone be sure that some parties will not try to improve their utility by ε? Once this possibility is allowed, it may snowball into opportunities for even greater gains by deviation. Indeed, Kol and Naor argue in favor of *strict* Nash equilibria, where parties will obtain strictly less utility by deviating (and show how to achieve strict equilibria in the presence of a simultaneous broadcast channel).

In our work, we achieve an exact notion of equilibrium (i.e. $\varepsilon = 0$). However, we allow a negligible probability that the honest parties will fail to compute the secret correctly, so our notion of "fairness" is also approximate. Nevertheless, we feel that the kind of error we achieve is preferable to ε-Nash. Indeed, the equilibrium concept is supposed to ensure that parties have an incentive to behave in a particular manner; if it is too weak, then parties may ignore it entirely and whatever analysis we do may be rendered irrelevant. On the other hand, if we achieve a sufficiently strong notion of exact equilibrium, then we may be confident that players will behave as predicted, and we are unlikely to see any bad events that are shown to occur with small probability under equilibrium play.

Trembling-Hand Equilibrium. In order to establish the equilibrium properties of the protocol, we introduce a framework of "extensive form games with public actions and private outputs," and use the formalism of incomplete information games to model players' uncertainties about the inputs (i.e. shares) of other players as well as uncertainty about which players are honest and which are rational. (For simplicity, we assume that each player is honest independently with some probability p, but with small modifications, the result should extend to other distributions on the set of honest players.) The solution concept of *Bayesian Nash equilibrium* handles the uncertainty that a player has about the shares dealt to other players and requires that beliefs are updated according to Bayes rule "whenever possible," meaning that this occurs when the observed actions are consistent with the equilibrium. A standard refinement is that of Bayesian *subgame perfect* Nash equilibrium, which captures the idea that the strategy is rational to follow regardless of the previous history of messages; intuitively, this means that the equilibrium does not rely on irrational empty threats (where a player will punish another player for deviating even at his own expense). In fact, we achieve the additional refinement of *trembling hand perfect equilibrium* [45], which strengthens this notion by requiring that players update their beliefs in a consistent and meaningful manner even when out-of-equilibrium play occurs. It is one of the strongest solution concepts studied for extensive form games, and was advocated in this context by Peter Bro Miltersen (personal communication) and Jonathan Katz [26].

Our Protocol. The protocol that we instruct honest players to follow is simple to describe. The participants take turns broadcasting their shares in sequence. However, if any of the first

[2]For round efficiency, sometimes people use a slightly more general channel where many parties can broadcast in a single round, but deviating parties are can perform 'rushing' — wait to see what others have broadcast before broadcasting their own values. We describe how to extend our results to this setting below.

$t-1$ parties deviates from the protocol by stopping and refusing to broadcast her share, then the protocol instructs all parties subsequent to the first $t-1$ to do the same. The intuition behind this protocol is that if there is likely to be at least one *honest* party after the first $t-1$ parties, then each rational party in the first $t-1$ parties will also have an incentive to reveal its share because by doing so, the honest party will also reveal her share and enable the rational parties to reconstruct the secret. Then we observe that as long as the set of honest parties is a random subset of $k = \omega(\log n)$ parties, and assuming that the total number, n, of players is sufficiently large, then there will be an honest party after party $t-1$ with all but negligible probability, as long as $t \leq (1 - \Omega(1)) \cdot n$. Thus, assuming that parties have a nonnegligible preference to learn the secret, we obtain an *exact* equilibrium in which *everyone* learns the secret with all but negligible probability.

In order to deal with the possibility that some players may try to reveal incorrect shares, we use information-theoretic message authentication codes (MACs) to authenticate the shares, following Kol and Naor [28]. Intuitively, we can tolerate the (negligible) forgery probability of the MACs (without getting an ε-Nash equilibrium) because the first $t-1$ players actually achieve *strictly* higher utility by revealing a valid share than by not doing so.

In addition, the incentives in our protocol hold regardless of what information the first $t-1$ players have about each others' actions, and similarly for the last $n-t+1$ players. Thus, our protocol can actually be implemented with only 2 rounds of communication (in contrast to all previous protocols, which required a super-constant number of rounds); we discuss how to formalize this below.

Modeling Contributions. While the intuition for our protocol is quite natural, modeling it game-theoretically turns out to be quite delicate. As discussed above, we introduce a Bayesian framework for capturing the uncertainty that players have about each others' secrets and which other players are honest vs. rational. Additional modeling contributions include:

Set Nash We find it useful to avoid specifying the exact actions that rational players should take in situations where the choice is irrelevant to the overall strategic and fairness properties of our protocol. We do this by developing a variant of the notions of Set-Nash [30] and CURB (Closed Under Rational Behavior) Sets [6] for extensive-form games and trembling-hand perfect equilibrium. Roughly speaking, this notion allows us to specify the equilibrium actions only in cases that we care about, and argue that players have no incentive to deviate from the specified actions provided that all other players are playing according to the specified actions (even if they may act arbitrarily when the action is unspecified) and given the existence of a small number of honest players. Since the honest strategy is consistent with the specified equilibrium actions, this solution concept ensures that even repeated rational deviations from the honest strategy (which we envision to be initial "program" distributed to all players) by all but a small number of players will keep everyone consistent with the specified actions. When this occurs as predicted, we show that all honest players will learn the secret with all but negligible probability, and thus fairness is maintained.

Modeling Rushing To save on rounds, the cryptography literature often allows protocols that specify messages for several players at once, but allows the possibility that deviating players may wait to see other players' messages before computing their own (i.e. simultaneity is allowed but not enforced). Modelling such "rushing" game-theoretically was posed as a challenge in the survey of Katz [26]. As mentioned above, we argue that our protocol can be collapsed to two rounds of communication. To capture the possibility of rushing game-theoretically, we follow an idea of Kalai [25], and argue that the specified strategy remains an

equilibrium for every ordering of players within each round. Thus, players have no incentive to wait for other players' messages; sending the same message will maximize their utility regardless of what other players send in the same round. (An alternative approach to modeling rushing would be to argue, in the spirit of an *ex post* Nash equilibrium, that players have no incentive to revise their messages after seeing the actions (and thus learning the private information) of others so long as the others play the equilibrium. We do not explore this possibility here.)

1.6 Future Directions and Independent Work

We view our work as but one more step in the line of work understanding the benefits of bringing together cryptography and algorithmic mechanism design. (See the survey [26].) While our main theorem is admittedly far from achieving an end goal that one would want to implement as is, we hope that our high-level message (regarding the benefit of a few honest players with many rational players) and our game-theoretic modelling (e.g. the Bayesian framework, the use of set-Nash, and the modelling of rushing) prove useful in subsequent work. Some specific ways in which our results could be improved are:

- Handling other distributions on (i.e. beliefs about) the set of honest players. Intuitively, this should be possible by having the dealer randomly permuting the order of the players and including the permutation in the shares (or publishing it). See Remark 5.1.
- Achieving solution concepts that are robust even to coalitional deviations from the protocol. In an earlier version of our paper [38], we demonstrated coalition-proofness (against "stable" coalitions) in a model that is even more simplified than the fail-stop one. As we have mentioned, Abraham et al. [2] show how to handle arbitrary, not necessarily stable, collusions of a small number of players.
- Generalizing from secret sharing to secure multiparty computation. Indeed, this is the main application for secret sharing and their reconstruction protocols.
- Getting stronger impossibility results for the entirely rational setting (prior impossibility results either require players to learn the secret with pprobability 1 [28], or suffered other restrictive constraints [39]) or, alternatively, finding a purely rational protocol.

O'Neill and Sangwan [37] extend the results from a preliminary version of our paper [39] in several ways, including achieving a strict Trembling Hand Perfect Equilibrium for a restricted deviation model (which is still more general than the fail-stop deviation model we considered in [39]) and handling a small number of "malicious players."

2 Definitions

2.1 Games with Public Actions and Private Outputs

To cast protocol executions into a game-theoretic setting, we introduce the notion of *extensive games with public actions and private outputs*. The basis of this new notion is the more standard definition of *extensive form games with perfect information*. Extensive form games enable us to model the *sequential* aspect of protocols, where each player considers his plan of action only following some of the other players' messages (the "*actions*" of the game-theoretic model). The perfect information

property captures the fact that each player, when making any decision in the public phase of the protocol, is perfectly informed of all the actions by other players that have previously occurred. Thus, extensive form games with perfect information are a good model for communication on a *synchronous broadcast channel*.

We build upon extensive form games with perfect information and augment them with an additional final *private* stage. This additional stage models the fact that at the end of the game, each player is allowed to take some arbitrary action as a function of the history $h \in H$ of messages so far. This action, along with the "history" of public actions that have taken place during the execution of the game (as well as the players' inputs) has a direct effect on players' payoffs.

Working in the framework of Bayesian games of incomplete information, players $i \in N$ are handed private inputs θ_i that belong to some pre-specified set Θ_i (the "*types*" of the game-theoretic model) and specify a distribution according to which the various inputs are chosen. Players' inputs can be thought of as the shares for the secret-sharing scheme, and are generated jointly with the actual secret. The secret is thought of as a "reference" value $\delta \in \Delta$ that is not given to the players at the beginning of the protocol (but may be determined by them through messages exchanged), and is used at the output stage along with private actions to determine player utilities; in game-theoretic terms, this secret can be considered to be picked by "Nature" in the first round of the game and then induces a distribution μ on the private types of agents. Nature sends each player a private message about its own type. Despite this uncertainty about other players' inputs, this remains a game of perfect information, with all previous actions of *other players* (just not that of Nature) known to each player. One additional change we make from the standard definition of Bayesian extensive-form games is that we allow the set of actions available to a player to depend on her type. This will be useful later on when we wish to model the possibility that some players are "honest," and will always follow a specified strategy even if it is not in her interest to do so. Whether a player is honest or rational will be indicated in her type, and honesty will be modeled by restricting the action sets of honest players to the message given by the specified strategy.

Definition 2.1 (Extensive game with public actions and private outputs)**.** *An* extensive form game with public actions and private outputs *is a tuple* $\Gamma = (N, H, P, A, L, \Delta, \Theta, \mu, u)$ *where*

- $N = \{1, \ldots, n\}$ *is a finite set of* players*,*
- Δ *is a (possibly infinite) set of possible* reference values*,*
- $\Theta = \Pi_{i \in N} \Theta_i$*, where* Θ_i *is a (possibly infinite) set of possible* private types *of player* $i \in N$*,*
- *values* $(\delta, \theta_1, \ldots, \theta_n) \in \Delta \times \Theta$*, are chosen jointly according to the distribution* μ*,*
- H *is a (possibly infinite) set of (finite)* history *sequences satisfying that the empty word* $\epsilon \in H$*. The components of a history sequence* $h \in H$ *are called* public actions*. A history* $h \in H$ *is* terminal *if* $\{a : (h, a) \in H\} = \emptyset$*. The set of terminal histories is denoted* Z*.*
- $P : (H \setminus Z) \to N$ *is a function that assigns a "next" player to every non-terminal history.*
- $A = (A_1, \ldots, A_n)$*, where* A_i *is a function that assigns for every pair* $(\theta_i, h) \in \Theta_i \times (H \setminus Z)$ *a finite set* $A_i(\theta_i, h)$ *of* public actions available to player $i = P(h)$*.*
- $L = (L_1, \ldots, L_n)$*, where* L_i *is a function that assigns for every pair* $(\theta_i, h) \in \Theta_i \times Z$ *a (possibly infinite) set* $L_i(\theta_i, h) \subseteq \Delta$ *of* private actions *available to player* $i \in N$*.*
- $u = (u_1, \ldots, u_n)$ *is a vector of payoff functions* $u_i : \Delta \times \Theta \times Z \times \Delta^n \to \mathbb{R}$*.*

In cases where Δ and H are finite, we say that Γ is finite.

An extensive form game with public actions and private outputs is interpreted as follows: the reference value and the types of the players are selected according to the joint distribution μ. The type $\theta_i \in \Theta_i$ is handed to player $i \in N$ and the value δ remains secret and affects the players' utilities. This is followed by a sequence of actions that are visible by all players. After any history $h \in H$, player $i = P(h)$ chooses an action from the set $A_i(\theta_i, h)$. The empty history $h_0 = (\epsilon)$ is the starting point of the game. Player $i_1 = P(h_0)$ chooses an action $a_1 \in A_{i_1}(\theta_{i_1}, h_0)$, which induces a history h_1. Then player $i_2 = P(h_1)$ subsequently chooses an action a_2 from the set $A_{i_2}(\theta_{i_2}, h_1)$; these choices determines the next actions of the players, and so on until a terminal history $h \in Z$ is reached. At this point, all players $i \in N$ simultaneously pick an action, b_i, from the set $L_i(\theta_i, h) \subseteq \Delta$, of available private actions. The utility (or payoff) of player i for an execution of the game is then determined to be the value $u_i(\delta, \theta, h, b_1, \ldots, b_n)$.

2.2 Strategies

The action chosen by a player for every history after which it is her turn to move, is determined by her *strategy* function. As is required in extensive-form games, the strategy is defined for all histories, even ones that would not be reached if the strategy is followed.

We refer to a pair $(\theta, h) \in \Theta \times H$ as a *state* of the game. A state (θ, h) is said to be *terminal* if $h \in Z$. Given our notion of extensive form games with public actions and private outputs, we distinguish between the *public strategy* of a player and her *private strategy*. The former is applied to (a player's view of) non-terminal states and is what determines a player's actions during the execution of the public part of the game. The latter is applied to (a player's view of) terminal states and is what determines a player's output. A strategy for player $i \in N$ is thus a pair $s_i = (m_i, f_i)$, where:

- The *public strategy* m_i is a function that takes a pair $(\theta_i, h) \in \Theta_i \times (H \setminus Z)$ and produces a public 'message', $m_i(\theta_i, h) \in A_i(\theta_i, h)$.
- The *private strategy* f_i is a function that takes a pair $(\theta_i, h) \in \Theta_i \times Z$ and produces a private 'output' $f_i(\theta_i, h) \in L_i(\theta_i, h)$.

A strategy s_i completely defines the actions of a player for every possible history $h \in H$ and every possible private input $\theta_i \in \Theta_i$. We will allow the public and private strategy functions to be mixed (i.e. randomized), where the randomization of the strategy is interpreted to be done independently at each application of the function, if a player has multiple moves in the game. Strategies that consist of deterministic functions are called *pure*, whereas strategies whose functions have full support on $A_i(\theta_i, h)$ are said to be *fully mixed*. We achieve fairness (with high probability) in a pure strategy equilibrium but use fully mixed strategies in defining the concept of trembling hand equilibrium.

We let $s = (s_1, \ldots, s_n)$ denote the vector of players' strategies, where $s_i = (m_i, f_i)$. The *outcome* o of the game Γ under strategy s is the random variable $(\delta, \theta, h, b_1, \ldots, b_n)$, where $(\delta, \theta) \in \Delta \times \Theta$ are sampled according to μ, $h \in H$ is the terminal state that results when each player $i \in N$ is given her type $\theta_i \in \Theta_i$, publicly follows the actions chosen by m_i, and computes her final private output b_i using f_i. That is, h is a history $h = (a_1, \ldots, a_\ell)$ such that for $j = 1, \ldots, \ell - 1$ and $i_{j+1} = P(a_1, \ldots, a_j)$, we have that $a_{j+1} \leftarrow m_{i_j}(\theta_{i_j}, (a_1, \ldots, a_j))$, and $b_i \leftarrow f_i(\theta_i, h)$ for all $i \in N$. The value of player i's utility is totally determined by the outcome o. The initial distribution, μ, of the secret and the shares, along with the strategies $s_i = (m_i, f_i)$ induce a distribution on the outcome o, and thus on the utilities of the players.

2.3 Nash and Set Nash Equilibria

Let $s = (m, f)$ be a strategy vector in Γ. Define $u_i(\mu, s)$ to be the expected value of the utility of player $i \in N$, when the types are selected according to the distribution μ and all players follow strategy s. (We consider games with histories with bounded length, and hence the expected value of the utility is well-defined.) We assume that rational players seek to maximize expected utility. The notion of Nash equilibrium (NE) captures the idea that no player should have an incentive to change her strategy, assuming that the other players play s_{-i}.

Definition 2.2 (Nash equilibrium)**.** *A strategy profile s is said to be a* (Bayesian) Nash equilibrium *for a game Γ if for all $i \in N$ and all strategies s'_i, $u_i(\mu, (s_{-i}, s'_i)) \leq u_i(\mu, s)$.*

In game theory, the notion of Bayesian NE is typically formulated to say that for each type θ_i of player i, player i maximizes her expected utility when the other players' types are distributed according to the conditional distribution $\mu|_{\theta_i}$. However, since a player's strategy depends on her type, this is equivalent to the formulation above, where we also take the expectation over θ_i.

The protocol that we introduce for secret reconstruction will be defined in a way that leaves the actions that a rational agent should follow only partially specified. This keeps the protocol as simple as possible and also minimizes the difference between the description of the behavior of an honest player and that of a rational player.

For this, we allow each player to have a strategy that maps each information set (i.e. view of the player) into a *set* of possible actions. More precisely, a *set-valued strategy* for player $i \in N$ is a pair $S_i = (M_i, F_i)$, where:

- The *public set-valued strategy* M_i is a function that takes a pair $(\theta_i, h) \in \Theta_i \times (H \setminus Z)$ and defines a set of public messages, $M_i(\theta_i, h) \subseteq A_i(\theta_i, h)$.
- The *private set-valued strategy* F_i is a function that takes a pair $(\theta_i, h) \in \Theta_i \times Z$ and defines a set of private outputs $F_i(\theta_i, h) \subseteq L_i(\theta_i, h)$.

We write $s_i = (m_i, f_i) \in S_i$ to indicate that strategy s_i is consistent with S_i, i.e. with $m_i(\theta_i, h) \in M_i(\theta_i, h)$ for all $(\theta_i, h) \in \Theta_i \times (H \setminus Z)$ and $f_i(\theta_i, h) \in F_i(\theta_i, h)$ for all $(\theta_i, h) \in \Theta_i \times Z$ (where these inclusions should hold with probability 1 in case s_i is a mixed strategy).

Definition 2.3 (Set Nash equilibrium)**.** *A profile $S = (S_1, \ldots, S_n)$ of set-valued strategies is a* (Bayesian) Set Nash equilibrium *for a game Γ if for all $i \in N$, every (possibly mixed) $s_{-i} \in S_{-i}$, there exists a strategy $s_i \in S_i$ so that for all strategies s'_i, $u_i(\mu, (s_{-i}, s'_i)) \leq u_i(\mu, s)$.*

Since all strategies s'_i for player i are weakly dominated by a strategy $s_i \in S_i$, when other players are playing strategy profiles consistent with S_{-i}, then there exists a Nash equilibrium that is consistent with S. To see this, consider a modified game in which all players are restricted to play consistently with S. This game has a Nash equilibrium s by Nash's existence theorem. Now, we argue that s must also be a Nash equilibrium even if players are not restricted to play consistently with S. If a player could strictly improve her utility by any deviation from s, then the definition of set-Nash guarantees that she could also do so playing consistently with S, contradicting the fact that s is a Nash equilibrium in the modified game.

A trival set-Nash equilibrium has S_i defined to include all possible strategies for all i. At another extreme, if S_i defines a singleton action set for all players at all information sets then this is a Nash equilibrium. Our goal will be to find set-Nash equilibrium for which every profile contained in it achieves some desirable property (namely everyone learning the secret). This allows us to avoid specifying moves that are not important for strategic properties.

Our definition of set-Nash equilibrium is stronger than the set-Nash equilibrium definition introduced by Lavi and Nisan [30], who require only that for every *pure* $s_{-i} \in S_{-i}$ there exists some $s_i \in S_i$ for which $u_i(\mu, s) \geq u_i(\mu, (s_{-i}, s'_i))$ for all possible strategies s'_i. This earlier definition of set-Nash is insufficient to ensure that there is a Nash equilibrium consistent with the set-valued strategy profile S. The problem is that upon restricting the game to S the only Nash equilibrium may be a mixed equilibrium, yet there may be some strategy s'_i outside of S_i that is strictly better than all $s_i \in S_i$ given that players $\neq i$ play a mixed strategy consistent with S_{-i}.

On the other hand, our definition of set-Nash is weaker than the CURB (Closed Under Rational Behavior) sets of Basu and Weibull [6]; see also a recent discussion in Benisch et al. [14]. A CURB set requires that for every mixed strategy s_{-i} consistent with S_{-i}, *all* best-responses for player i are consistent with set-valued strategy S_i whereas we require only that there is *some* best-response that is consistent with S_i. Given our set-Nash construct, it is rational for a rational player to play inside the set given that other rational payers play inside the set. The strengthening provided by CURB allows one to further argue that rational players will not play outside of the set, and provides the tightest set-valued approximation to Nash equilibrium in the sense of containing *all* Nash equilibrium solutions.

2.4 Trembling-Hand Perfect Equilibrium

The basic notion of Nash equilibrium (or even strict Nash equilibrium) turns out to be an unsatisfactory solution concept for extensive-form games. The reason is that a Nash equilibrium can rely on "incredible" threats by players — ones that are needed to maintain the equilibrium but never occur during the equilibrium play and would not be in the self-interest of the player if tested. A more appealing solution concept is that of *subgame perfect equilibrium*. This is a standard strengthening of the notion of Nash equilibrium in that it requires that the equilibrium strategy is a Nash equilibrium in every *subgame* of the original extensive form game.

In order to adapt the subgame perfect equilibrium concept to the Bayesian setting one has to take into consideration the inherent uncertainty about player's types (and as a result about the actions implied by their strategies). The assumption underlying the Bayesian setting is that individual players have *beliefs* about the values of other players' types. The beliefs are in fact distributions from which players think that the types of other players were drawn. At the beginning of the game, the belief corresponds to the initial distribution μ conditioned on the player's knowledge of her own type. As this game of incomplete information progresses, players update their beliefs as a function of other players' actions (recall that a player's actions may depend on her type).

At the heart of a solution concept for extensive-form games with incomplete information is a requirement about the way in which the players update their beliefs. A straightforward approach for a player to update her beliefs is to use Bayes rule to condition on her own view of the actions taken in the game (as represented in her "information set"). This is the basic approach taken in the game theory literature, and the one pursued in a previous version of this paper [39]. But such an approach suffers from the drawback that updating is not well-defined for views that occur with zero probability, i.e. following out of equilibrium play. See [19] for further discussion of the difficulties related to Bayesian updating in extensive form games.

A stronger approach, also discussed in the game theory literature, is the one of *trembling-hand perfect equilibrium* [45]. The idea behind trembling-hand perfect equilibrium is that updating is not problematic if the strategies under consideration are fully mixed (since such strategies would never incur an updating that conditions on a zero probability event). It thus becomes natural to require that the equilibrium strategy is a best response in every subgame to some sequence of fully

mixed strategies that converge to equilibrium, and this is indeed the definition of trembling hand equilibria. An interpretation of trembling-hand perfect equilibrium is that it requires each player's strategy to be robust to beliefs that are generated by play that could be the result of small mistakes.

A trembling-hand perfect equilibrium is also a Bayesian subgame perfect equilibrium, and when applied to normal form games has the useful property of eliminating weakly dominated strategies.[3] Trembling-hand perfect equilibrium is one of the strongest solution concepts studied for extensive form games, and an extensive-form trembling-hand perfect equilibrium is also a sequential equilibrium (due to Kreps and Wilson [29]).

The trembling-hand solution concept builds on the following definition of a subgame, which captures the way in which players update their beliefs as a result of other players' actions.

Definition 2.4 (Subgame)**.** *Let Γ be an extensive form game with public actions and private outputs and $s = (m, f)$ a* fully mixed *strategy profile for Γ. For a history $h \in H$ the* subgame $\Gamma(s, h) = (N, H|_h, P|_h, A|_h, L|_h, \Delta, \Theta, \mu|_h, u|_h)$ *and* substrategy $s|_h$ *are defined as follows.*

- *$H|_h$ is the set of sequences h' for which: $(h, h') \in H$; i.e., $H|_h = \{h : (h, h') \in H\}$.*
- *For every $h' \in H|_h$, the function $P|_h$ is defined by $P|_h(h') = P(h, h')$.*
- *For every $i \in N$, the function $A_i|_h$ is defined by $A_i|_h(\theta_i, h') = A_i(\theta_i, (h, h'))$.*
- *For every $i \in N$, the function $L_i|_h$ is defined by $L_i|_h(\theta_i, h') = L_i(\theta_i, (h, h'))$.*
- *For every $i \in N$, the function $u_i|_h$ is defined by $u_i|_h(\delta, \theta, h', b) = u_i(\delta, \theta, (h, h'), b)$.*
- *The distribution $\mu|_h$ is defined as μ conditioned on h given strategy s.*[4]
- *The substrategy $s|_h = (m|_h, f|_h)$ is defined by $m_i|_h(\theta_i, h') = m_i(\theta_i, (h, h'))$ and $f_i|_h(\theta_i, h') = f_i(\theta_i, (h, h'))$.*

Given a strategy profile s and a strategy s'_i for player i, we denote by $u_i|_h(\mu|_h, s_{-i}|_h, s'_i|_h)$ the expected value of player i's utility under strategy vector $(s_{-i}|_h, s'_i|_h)$ in the game $\Gamma(s, h)$. This is interpreted as considering player i's expected utility when all players except player i follow strategies s_{-i}, and assuming that until the history h has been reached player i has played according to strategy s_i, and from that point on according to strategy s'_i.

Definition 2.5 (Trembling-hand perfect equilibrium)**.** *Let $\Gamma = (N, H, P, A, L, \Delta, \Theta, \mu, u)$ be an extensive form game with public actions and private outputs. A strategy profile s is said to be a* trembling-hand perfect equilibrium *for Γ if there exists a sequence of fully mixed strategy profiles $(s^k)_{k=0}^{\infty}$ converging to s so that for every history $h \in H$, every $i \in N$, and every strategy s'_i for player i in the game $\Gamma(s^k, h)$ it holds that $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i|_h) \geq u_i|_h(\mu|_h, s^k_{-i}|_h, s'_i)$ in the game $\Gamma(s^k, h)$ for all k.*

Our definition of a trembling-hand perfect equilibrium requires that player i's strategy s_i defines a best-response at every partial view/information set (because the strategies are fully mixed), where past perturbations of its own strategy are also considered within s^k in defining subgame $\Gamma(s^k, h)$. This definition is slightly weaker than the standard trembling-hand definition for extensive-form games in that it does not allow future trembles for player i (since we are replacing all of player i's

[3] See Osborne and Rubinstein [40] for an extended discussion, including an example to show that trembling hand perfect equilibria of extensive form games can still include weakly dominated strategies.

[4] Since s is fully mixed then every h occurs with nonzero probability, and so conditioning on h is well defined.

future actions in s^k with the non-trembling s_i).[5] We adopt this definition because it is conceptually simpler, but our protocol actually achieves (a set Nash analogue of) the standard definition.

We define a set-Nash analogue of trembling-hand perfect equilibrium. To the best of our knowledge, such a combination has not been previously considered in the literature. Given a history $h \in H$ and a set valued strategy S_i, we define $S_i|_h$ in the natural way (i.e., if $S_i = (M_i, F_i)$ then $S_i|_h = (M_i|_h, F_i|_h)$ where $M_i|_h(\theta_i, h') = M_i(\theta_i, (h, h'))$ and $F_i|_h(\theta_i, h') = F_i(\theta_i, (h, h'))$).

Definition 2.6 (Trembling-hand perfect set equilibrium)**.** *Let $\Gamma = (N, H, P, A, L, \Delta, \Theta, \mu, u)$ be an extensive form game with public actions and private outputs. A profile $S = (S_1, \ldots, S_n)$ of set-valued strategies is said to be a* trembling-hand perfect set equilibrium *for Γ if for every $s \in S_1 \times \cdots \times S_n$ there exists a sequence of fully mixed strategy profiles $(s^k)_{k=0}^{\infty}$ converging to s so that for every history $h \in H$ and every $i \in N$, there exists a strategy $s_i' \in S_i|_h$ such that for all strategies s_i'' in the game $\Gamma(s^k, h)$ it holds that $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i') \geq u_i|_h(\mu|_h, s^k_{-i}|_h, s_i'')$ in the game $\Gamma(s^k, h)$ for all k.*

There is no requirement in the definition for the sequence of perturbed strategies to be consistent with the profile S of set-valued strategies.

Lemma 2.7 (One-deviation property trembling-hand perfect set equilibrium)**.** *Let Γ be a finite extensive form game with public actions and private outputs and histories with bounded length. The set-valued strategy profile S is a trembling-hand perfect set equilibrium if and only if for every $s \in S$, there exists a fully mixed sequence $(s^k)_{k=0}^{\infty}$ converging to s, so that for every player $i \in N$, every pair $(\theta_i, h) \in \Theta_i \times H$ such that $P(h) = i$, and every strategy s_i'' for player i in the game $\Gamma(s^k, h)$ such that $s_i''(\theta_i, \epsilon) \notin S_i|_h(\theta_i, \epsilon)$, there is a strategy s_i' for player i in game $\Gamma(s^k, h)$ such that $s_i'(\theta_i, \epsilon) \in S_i|_h(\theta_i, \epsilon)$ and $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i') \geq u_i|_h(\mu|_h, s^k_{-i}|_h, s_i'')$ in game $\Gamma(s^k, h)$ for all k*

Proof. It is easy to see that this one-deviation property is necessary for a trembling-hand perfect set equilibrium. For the sufficient direction, fix game Γ, some strategy profile $s \in S$ in set-valued strategy profile S, and consider a fully mixed sequence $(s^k)_{k=0}^{\infty}$ converging to s for which the one-deviation property holds. Proceed by backwards induction on the number of rounds in the subgame $\Gamma(s^k, h)$. The induction hypothesis is that for every history $h \in H$ for which game $\Gamma(s^k, h)$ has $z \geq 1$ or less rounds, and every player $i \in N$, there exists a strategy $s_i' \in S_i|_h$ such that for all strategies s_i'' in the game $\Gamma(s^k, h)$ it holds that $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i') \geq u_i|_h(\mu|_h, s^k_{-i}|_h, s_i'')$ for all k.

For the base case, game $\Gamma(s^k, h)$ has one round. Fix player i. If $P(h) \neq i$ then $s_i' \in S_i|_h$ trivially satisfies $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i') \geq u_i|_h(\mu|_h, s^k_{-i}|_h, s_i'')$ for all s_i'' in the subgame. If $P(h) = i$ (the interesting case) then by the one-deviation property, for any $s_i'' \notin S_i|_h$ there exists a strategy s_i' in the subgame with at least as much utility and such that $s_i'(\theta_i, \epsilon) \in S_i|_h(\theta_i, \epsilon)$, and thus $s_i' \in S_i|_h$ since the subgame has one round. To establish the inductive hypothesis, pick as a best-response the strategy $s_i' \in S_i|_h$ with highest utility $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i')$, when assuming that other players play according to s^k_{-i}.

For the inductive case, consider a game $\Gamma(s^k, h)$ with $z > 1$ rounds. Fix player i. Again, the case with $P(h) \neq i$ can be trivially established from the inductive hypothesis on subgames with $z - 1$ rounds. Consider the interesting case of $P(h) = i$, and fix s_i'' in game $\Gamma(s^k, h)$ such that $s_i''(\theta_i, \epsilon) \notin S_i|_h(\theta_i, \epsilon)$. By the one-deviation property, there exists a s_i' in game $\Gamma(s^k, h)$ such that $s_i'(\theta_i, \epsilon) \in S_i|_h(\theta_i, \epsilon)$ with at least as much utility as s_i''. Pick as a candidate best-response to s^k_{-i} the strategy $s_i' \in S_i|_h$ with the highest utility $u_i|_h(\mu|_h, s^k_{-i}|_h, s_i')$: this strategy s_i' will have at least as much utility as any s_i'' in the subgame. Now appeal to the inductive hypothesis and modify s_i'

[5] The considerations that require that a player also consider its own future trembles can be modeled via the "agent strategic form" and are relevant in the case that there are multiple public plays by a player. See Osborne and Rubinstein [40] for an extended discussion.

on all subsequent subgames to adopt a strategy $\tilde{s}_i$ consistent with S_i, and with at least as much utility. This modified strategy establishes the inductive hypothesis and completes the proof. ■

The differences between this formulation and Def 2.6 are: (1) we switch order of quantifiers, saying for every strategy s_i'', there is something at least as good consistent with S_i (rather than requiring some s_i' consistent with S_i that is at least as good as every s_i''); and (2) we do not require that s_i' is consistent with S_i everywhere, but only at the history h. This is the sense in which it is a one-deviation-like property (but note that we do allow s_i'' to differ from s in many places). Later deviations from S_i are implicitly dealt with when considering longer histories. This one-deviation property provides a useful simplification to the analysis of the secret-sharing game.

3 Secret Sharing

In this section we formally define the idea of a secret sharing scheme. We then define what is a *secret sharing reconstruction protocol*, and cast it into a game-theoretic setting by defining a corresponding *reconstruction game* (as induced by a given reconstruction protocol). The latter allows us to reason about whether players have an incentive to follow the strategy specified by the protocol, even if they are allowed arbitrary deviation from the protocol's instructions (including a change of their private output).

3.1 Secret-Sharing Schemes

A secret sharing scheme describes the basic requirements of secret sharing and is not game-theoretic. In order to define a secret sharing scheme we need to specify the sets from which the secret and its shares are drawn, as well as a joint probability distribution under which the shares of a secret are generated by the dealer (along with the secret).

Definition 3.1 (Secret sharing)**.** *A* threshold secret sharing scheme *is a tuple* $(N, t, \Delta, \Theta, \mu, g)$ *where*

- $N = \{1, 2, \ldots, n\}$ *is the set of* players*,*
- $t \in \mathbb{N}$ *is the* threshold *of the scheme,*
- Δ *is the set from which the "secret" is chosen,*
- $\Theta = \Pi_{i \in N} \Theta_i$*, where* Θ_i *is the set of possible shares of player* $i \in N$*,*
- μ *is a joint probability distribution for the secret and the shares* $(\delta, \theta_1, \ldots, \theta_n) \in \Delta \times \Theta$*,*
- $g = \{g_S : \prod_{i \in S} \Theta_i \to \Delta \cup \{\lambda\}\}_S$ *is a collection of* reconstruction functions*, such that:*
  - *for each set* $S \subseteq N$ *of size strictly less than* t*,* $g_S((\theta_i)_{i \in S}) = \lambda$*,*
  - *for each set* $S \subseteq N$ *of size at least* t*:*

    $$\Pr[g_S((\theta_i)_{i \in S}) = \delta] = 1$$

    where $(\delta, \theta_1, ..., \theta_n)$ *are drawn according to* μ*.*
  - *for every* $S \subseteq N$ *and* $(\theta_i)_{i \in S}$ *such that* $g_S((\theta_i)_{i \in S} \neq \lambda$ *it holds that*

    $$g_T((\theta_i)_{i \in T}) = g_S((\theta_i)_{i \in S})$$

    for every $T \subseteq S$ *of size at least* t*, and*

- *for every $S \subseteq N$ of size less than t, the tuple $(\delta, (\theta_i)_{i\in S})$ has the same distribution as $(U, (\theta_i)_{i\in S})$ when δ and the θ_i's are chosen according to μ and U is uniformly and independently chosen in Δ,*

A secret sharing scheme is implemented by letting a trusted *dealer* jointly pick the secret and shares according to the distribution μ, and then distributing share $\theta_i \in \Theta_i$ to player $i \in N$. The reconstruction functions are what enables any set S of at least t players to use their shares $(\theta_i)_{i\in S}$ in order to jointly reconstruct the secret. The scheme should also guarantee secrecy against any subset S of less than t players. This requirement is expressed in the last item of Definition 3.1.

Our analysis will require the secret sharing to satisfy some additional properties. The properties we require are natural in the context of secret sharing, are satisfied, for example, by the Shamir secret sharing scheme [46].

Definition 3.2 (Admissible secret sharing)**.** *A threshold secret sharing scheme $(N, t, \Delta, \Theta, \mu, g)$ is said to be* admissible *if it satisfies the following conditions:*

- *for every $S \subseteq N$ of size less than t, the tuple $((\theta_i)_{i\in S})$ is uniformly distributed in $\prod_{i\in S} \Theta_i$ and independent from δ, when δ and the θ_i's are chosen according to μ.*
- *for every $S \subseteq N$ of size t, any $j \notin S$, and every $\theta_j \in \Theta_j$ so that $g_{S\cup\{j\}}((\theta_i)_{i\in S}, \theta_j) \neq \lambda$, there exists a unique $\delta \in \Delta$ so that $g_{S\cup\{j\}}((\theta_i)_{i\in S}, \theta_j) = \delta$.*

To prevent players from revealing shares that are different than the ones they were dealt, we will want to work with a secret sharing scheme that is *authenticated*.

Definition 3.3 (Authentication Scheme)**.** *An* authentication scheme *with* forgery probability $\gamma \in [0, 1]$ *is a tuple* $(M, T, K, \nu, \mathrm{Auth}, \mathrm{Ver})$ *such that:*

- *M is a set of* messages*, T is a set of* tags*, and K is a set of* keys*,*
- *ν is a probability distribution on K,*
- $\mathrm{Auth} : M \times K \to T$ *and* $\mathrm{Ver} : M \times T \times K \to \{0, 1\}$*, are functions so that for every $m \in M$ and for κ that is chosen according to ν:*
  - $\mathrm{Ver}(m, \tau, \kappa) = 1$ *with probability 1 where* $\tau = \mathrm{Auth}_\kappa(m) = \mathrm{Auth}(m, \kappa)$*,*
  - *For every function $A : M\times T \to M\times T$, the probability that* $\mathrm{Ver}(A(m, \mathrm{Auth}_\kappa(m, \tau)), \kappa) = 1$*, and the first output of $A(m, \tau)$ is not equal to m is at most γ.*

A standard technique for authenticating shares in a (plain) secret sharing scheme would be the following [49, 44, 28]: the dealer chooses at random elements $\alpha_i, \beta_i \in \mathbb{F}, \beta_i \neq 0$, and sets $\tau_i = \alpha_i \cdot \theta_i + \beta_i \in \mathbb{F}$, where $\theta_i \in \mathbb{F}$ is player i's share (in particular it is required that $|\mathbb{F}| > |\Delta|$). The value τ_i (the tag) is given to player i, along with the share θ_i. The other players each get the pair $\kappa_i = (\alpha_i, \beta_i)$ (the key). (Thus, in this example $K = \mathbb{F} \times (\mathbb{F} \setminus \{0\})$ and $T = \mathbb{F}$.)

In order to prove that θ_j is her true share, player j is required to broadcast the tag τ_j. The other players can then verify with high probability by checking that $\tau_j = \alpha_j \cdot \theta_j + \beta_j$ (the probability of forgery can be made negligibly small by picking a sufficiently large $\mathbb{F}$). Thus, in effect, the share dealt by the dealer consists of a tuple $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j\neq i})$, where $\tau_i = \alpha_i \cdot \theta_i + \beta_i \in \mathbb{F}$, $\kappa_j = (\alpha_j, \beta_j)$ and $\theta_i \in \mathbb{F}$ is the original share chosen by the dealer. The distribution according to which the dealer chooses the authenticated shares is thus $\tilde{\mu} = \mu \times \nu^{|N|}$. The above technique is applicable to any secret sharing scheme (given an appropriate embedding of elements in Δ into elements in $\mathbb{F}$).

3.2 Reconstruction Protocols

Once shares are distributed among the players, it is required to specify a protocol according to which the players can jointly reconstruct the secret at a later stage (using the reconstruction function). The reconstruction protocol prescribes a way in which the players compute their "messages", which are chosen from a given fixed "alphabet," and are then broadcast to all other players. The protocol also specifies an *output function* that is used by the players to compute their (private) output.

One specific way of doing so would be to let the players broadcast their private shares in some order. However, to avoid restrictions on the way in which the protocol proceeds, we give a more general definition that allows the exchange of arbitrary messages.

Definition 3.4 (Secret sharing reconstruction protocol)**.** *A* reconstruction protocol *for a secret sharing scheme* $(N, t, \Delta, \tilde{\Theta}, \tilde{\mu}, g)$ *is a tuple* $\Pi = (\Sigma, H, P, m^*, f^*)$ *where*

- Σ *is a finite set of* messages*,*
- $H \subseteq \Sigma^*$ *is a (possibly infinite) set of* protocol history *sequences, satisfying that the empty word* $\epsilon \in H$*. We let* $M(h) = \{m : (h, m) \in H\} \subseteq \Sigma$*. A history* $h \in H$ *is* terminal *if* $M(h) = \emptyset$*. The set of terminal histories is denoted* Z*.*
- $P : (H \setminus Z) \to N$ *assigns a set of "next" players to every non-terminal history.*
- $m^* = (m_1^*, \ldots, m_n^*)$ *is a vector of* next-message functions*, where the function* m_i^* *maps a pair* $(\tilde{\theta}_i, h) \in \tilde{\Theta}_i \times (H \setminus Z)$ *to a message,* $m_i^*(\tilde{\theta}_i, h) \in \Sigma$*, for every history* h *such that* $i = P(h)$*.*
- $f^* = (f_1^*, \ldots, f_n^*)$ *is a vector of* output functions $f_i^* : \tilde{\Theta}_i \times Z \to \Delta$*.*

A reconstruction protocol for a given secret sharing scheme is implemented under the assumption that the secret and shares $(\delta, \theta_1, \ldots, \theta_n)$ are chosen according to the distribution μ, and $(\tau_i, \kappa_i)_{i \in N}$ are chosen according to ν. Player i's type is $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j \neq i})$. The protocol is interpreted as follows: player $i = P(h)$ chooses a message $m = m_i^*(\tilde{\theta}_i, h) \in \Sigma$. The empty history $h_0 = \epsilon$ is the starting point of the execution. Player $i_0 = P(\epsilon)$ chooses a message $m = m_{i_0}^*(\tilde{\theta}_{i_0}, \epsilon) \in \Sigma$. This induces a history $h_1 = (m)$, where $\bar{m} = (m_1)$, and player $P(h_1)$ subsequently chooses a message from the set Σ; this choice determines the next player to move, and so on until a terminal history $h \in Z$ is reached. At this point all players can determine the value of their private output functions, $f_i^*(\tilde{\theta}_i, h)$. Generally, we are interested in secret-sharing protocols in which all players will compute the secret correctly (i.e. $f_i^*(\tilde{\theta}_i, h) = \delta$ with high probability over μ, provided all players follow the protocol). Rather than require this as part of the definition, however, we will address explicitly this in the statements of our positive and negative results.

3.3 Reconstruction Games

Based on the definition of a secret sharing protocol, we may now formalize an induced *reconstruction game*. Loosely speaking, this is an interpretation of a reconstruction protocol as an extensive form game with public messages and private outputs, in which arbitrary deviations from the protocol's instructions are allowed. The interpretation of the protocol as a game is straightforward: protocol histories correspond to game histories, messages in the protocol correspond to actions, next message functions correspond to strategies, and the outputs correspond to output actions.

The reconstruction game allows player $i \in N$ the choice between continuing with the protocol's prescribed instructions (and in particular choosing an action according to m_i^*), and deviating from Π (by sending some other message from Σ).

Definition 3.5 (Reconstruction game for secret sharing)**.** *A* reconstruction game *that corresponds to reconstruction protocol* $\Pi = (\Sigma, H, m^*, f^*)$ *for authenticated secret sharing scheme* $(N, t, \Delta, \tilde{\Theta}, \tilde{\mu}, g)$ *is an extensive form game with public actions and private outputs* $\Gamma = (N, H, P, A, L, \Delta, \tilde{\Theta}, \tilde{\mu}, u)$ *satisfying the following conditions:*

- *The set of private actions available to player* $i \in N$ *is* $L_i = \Delta$*.*
- *For every nonterminal state* $(\tilde{\theta}_i, h) \in \tilde{\Theta}_i \times (H \setminus Z)$*, the set of available* public actions *to player* $i = P(h)$*, is* $A_i(\tilde{\theta}_i, h) = \Sigma$*.*
- *For an outcome* $o = (\delta, \tilde{\theta}, h, b_1, \ldots, b_n)$*, the utilities* $u_i(o)$ *are a function of only* i *and the set* $S(o) = \{j : b_j = \delta\}$*. Moreover, we require that:*
  1. *If* $i \in S(o)$ *and* $i \notin S(o')$*, then* $u_i(o) > u_i(o')$*,*
  2. *If* $S(o) \subsetneq S(o')$ *and either* $i \in S(o) \cap S(o')$ *or* $i \notin S(o) \cup S(o')$*, then* $u_i(o) > u_i(o')$*.*

The honest strategy vector *in* Γ *is the pair* $s^* = (m^*, f^*)$*.*

A reconstruction game is interpreted as follows: a dealer selects a secret, shares and authentication information from the distribution $\tilde{\mu}$ and hands the input $\tilde{\theta}_i \in \Theta_i$ to player $i \in N$. This is followed by a sequence of actions (messages) that are prescribed by the reconstruction protocol Π, and are in particular visible by all players. At each point $h \in H$ of the game, player $i = P(h)$ faces a decision of whether to continue according to the prescribed strategy m_i^*, or to deviate from the prescribed instruction. In case the player has chosen to follow the strategy, she broadcasts the message $m_i^*(\tilde{\theta}_i, h)$. Otherwise, she may choose an arbitrary message from the set Σ. Once a terminal state $(\tilde{\theta}, h)$ is reached, we allow each player to choose any private output from $L_i = \Delta$.

The utility function is defined so that a player first prefers to learn the secret over not learning the secret, and secondly (all other things being equal) prefers that fewer other players learn the secret. It will be convenient for us to restrict attention to utility functions that are *linear* in the following sense:

Definition 3.6. *A reconstruction game as above has* linear utilities *if there are real numbers* $(\alpha_{ij})_{i,j \in N}$ *such that for every outcome* $o = (\delta, \theta, h, b_1, \ldots, b_n)$ *and player* $i \in N$*, we have* $u_i(o) = \sum_{j \in N} \alpha_{ij} \cdot I(b_j = \delta)$*, where* I *denotes an indicator function. We refer to ratio* $\rho = \min_i(\alpha_{ii} / -\sum_{j \neq i} \alpha_{ij})$ *as the players'* preference for learning the secret*.*

The constraints on the utility functions in Def. 3.5 imply that $\alpha_{ii} > 0$ for all i (players want to learn the secret), $\alpha_{ij} < 0$ for $i \neq j$ (players want others not to learn the secret), and $\sum_j \alpha_{ij} > 0$ for all i (players want to learn the secret more than they want others not to learn); hence $\rho > 1$.

4 Our Protocol

4.1 Introducing an honest minority

Our goal is to show that every authenticated secret-sharing scheme has a reconstruction protocol so that any reconstruction game that corresponds to it has an equilibrium strategy in which all players learn the secret. To do this, we require that a small subset of *honest* players in the reconstruction game always follows the strategy prescribed by the reconstruction protocol (whether or not this is the best response to other players' actions). We model this scenario by assuming that the set of honest players is selected according to some distribution that will specify to each player whether

she is to act honestly or rationally. The set of actions of an honest player will be then restricted to coincide with the strategy prescribed by the reconstruction protocol. The set of actions of a rational player will remain unchanged (i.e., as in the original definition of a reconstruction game).

Definition 4.1 (Reconstruction game with honest players)**.** *Let* $\Pi = (\Sigma, H, m^*, f^*)$ *be a reconstruction protocol for an authenticated secret sharing scheme* $(N, t, \Delta, \tilde{\Theta}, \tilde{\mu}, g)$*. A* reconstruction game that corresponds to Π with honest players *is an extensive form game with public actions and private outputs* $\Gamma = (N, H, P, A, L, \Delta, \tilde{\Theta} \times \Omega, \tilde{\mu} \times \zeta, u)$*, satisfying the following conditions:*

- $\Omega = \Pi_{i \in N} \Omega_i$*, where* $\Omega_i = \{\texttt{honest}, \texttt{rational}\}$ *indicates whether player* $i \in N$ *is honest or rational, and* $\tilde{\Theta}_i \times \Omega_i$ *is the set of possible types for player* $i \in N$*,*
- ζ *is a distribution on* Ω*, and values* $(\delta, \tilde{\theta}, \omega) = (\delta, \tilde{\theta}_1, \ldots, \tilde{\theta}_n, \omega_1, \ldots, \omega_n) \in \Delta \times \tilde{\Theta} \times \Omega$ *are chosen according to the distribution* $\tilde{\mu} \times \zeta$*. We refer to* ζ *as the* honest-player distribution*.*
- *For every nonterminal state* $(\tilde{\theta}, h) \in \tilde{\Theta} \times H$ *the set of available* public actions *to player* i*, is*

$$A_i(\tilde{\theta}_i, \omega_i, h) = \begin{cases} \{m_i^*(\tilde{\theta}_i, h)\} & \text{if } w_i = \texttt{honest} \\ \Sigma & \text{if } w_i = \texttt{rational} \end{cases}$$

- *For every terminal state* $(\tilde{\theta}, h) \in \tilde{\Theta} \times Z$ *the set of available* private actions *to player* i*, is*

$$L_i(\tilde{\theta}_i, \omega_i, h) = \begin{cases} \{f_i^*(\tilde{\theta}_i, h)\} & \text{if } w_i = \texttt{honest} \\ \Delta & \text{if } w_i = \texttt{rational} \end{cases}$$

- *For an outcome* $o = (\delta, (\tilde{\theta}, \omega), h, b_1, \ldots, b_n)$*, the utilities* $u_i(o)$ *are a function of only* i *and the set* $S(o) = \{j : b_j = \delta\}$*. Moreover, we require that:*
  1. *If* $i \in S(o)$ *and* $i \notin S(o')$*, then* $u_i(o) > u_i(o')$*,*
  2. *If* $S(o) \subsetneq S(o')$ *and either* $i \in S(o) \cap S(o')$ *or* $i \notin S(o) \cup S(o')$*, then* $u_i(o) > u_i(o')$*.*

The honest strategy vector *in* Γ *is the pair* $s^* = (m^*, f^*)$*.*

We interpret a reconstruction game with honest players as follows. The private type of player $i \in N$ consists of a pair $(\tilde{\theta}_i, \omega_i) \in \tilde{\Theta}_i \times \Omega_i$ that is drawn along with other player's types and the reference value δ according to the distribution $\tilde{\mu} \times \zeta$. The value of $\omega_i \in \{\texttt{honest}, \texttt{rational}\}$ determines whether player i is bound to follow the honest strategy (as prescribed by Π), or will be allowed to deviate from it. The constraints on the set of actions of each player create a situation in which rational players are indeed free to deviate from the public strategy vector m^* (since they are allowed to choose any action in Σ), whereas the honest players are in fact restricted to the single action prescribed by m^*.

4.2 Our Main Result

We will show that assuming the existence of a small number of honest players, there is a reconstruction protocol such that every corresponding reconstruction game has an equilibrium such that with high probability all players learn the secret with certainty, provided that the set of honest players is uniform among all sets of a sufficiently large size and every player has a nonnegligible preference for learning the secret. Specifically, our theorem is the following:

Theorem 4.2. *Every authenticated secret-sharing scheme $(N, t, \Delta, \tilde{\Theta}, \tilde{\mu}, g)$, with $t < |N| = n$, has a reconstruction protocol Π such that the following holds. Let $\Gamma = (N, H, P, A, L, \Delta, \tilde{\Theta} \times \Omega, \tilde{\mu} \times \zeta_m, u)$ be a reconstruction game that corresponds to Π with honest players and linear utility functions, where ζ_m is a distribution over tuples $(\omega_1, \ldots, \omega_n) \in \Omega$ for which $\omega_i =$ `honest` with probability m/n independently for all $i \in N$, for some real number $m \in [0, n]$. Suppose further that the players' preference ρ for learning the secret satisfies:*

$$\rho > \frac{1 - 1/|\Delta|}{1 - 1/|\Delta| - p(n, m) - \gamma} \tag{1}$$

where $p(n, m) = (1 - m/n)^{n-t+1} \leq \exp(-m \cdot (n - t)/n)$ and γ is the forgery probability of the authenticated secret-sharing scheme. Then Γ has a profile $S = (S_1, \ldots, S_n)$ of set-valued "rational" strategies such that:

1. *The honest strategy profile $s^* = (m^*, f^*)$ is consistent with S,*
2. *S is a trembling-hand perfect set equilibrium in Γ,*
3. *For any strategy vector $s \in S$, the probability that all* honest *players compute the secret correctly in Γ is at least $1 - (n - t + 1)\gamma - p(n, m)$, when the players' types are chosen according to μ and they follow strategy vector s.*
4. *For any trembling-hand perfect Nash equilibrium $s \in S$, the probability that* all *players compute the secret correctly in Γ is at least $1 - (n - t + 1)\gamma - p(n, m)$, when the players' types are chosen according to μ and they follow strategy vector s.*
5. *S does not depend on the particular utility functions u in Γ (provided they satisfy (1)).*

In the common case that $t \leq (1 - \Omega(1)) \cdot n$, observe that $p(n, m) = \exp(-\Omega(m))$ is negligible provided that $m = \omega(\log n)$, i.e. the expected number of honest players is superlogarithmic. If, in addition, the forgery probability γ is negligible then Condition 1 simply says that a player's preference for learning the secret should not be negligible.

The honest-player distribution ζ_m models a situation in which players have no information about which other players are honest or rational, except for some a priori belief on the probability a given player is honest. With small modifications, the theorem should extend to other distributions ζ as well; see Remark 5.1.

4.3 The reconstruction protocol

Let $(N, t, \Delta, \tilde{\Theta}, \tilde{\mu}, g)$ be an authenticated secret-sharing scheme with $t < |N|$ and $|\Delta| > 1$. We assume that a dealer distributes shares to the players according to the distribution μ, and would like to design a protocol $\Pi = (\Sigma, H, m^*, f^*)$ for secret share reconstruction.

The protocol proceeds in two stages, where in the first stage a subset of $t-1$ players is instructed to reveal their share to all other parties in some sequence, and in the second stage the remaining $n - t + 1$ players are instructed to reveal their share in some sequence, provided that none of the $t - 1$ players in the first stage has failed to reveal her share. The individual parties will reveal their share using a *synchronous broadcast* channel (as modeled by our definitions of reconstruction protocols (Def. 3.4) and extensive games with public actions and private outputs (Def. 2.1)).

The stage in which a player is instructed to broadcast is fixed in some arbitrary manner. For the sake of concreteness, suppose that at stage 1 of the protocol, it is the turn of players $1, \ldots, t-1$

to broadcast, and that at stage 2 it is the turn of players $t, \ldots, n$. The protocol will instruct player i to either *reveal* her share $\theta_i \in \Theta_i$ or not to reveal anything (symbolized by a special action denoted $\perp \in \Sigma$). Specifically, we will require that player i reveals θ_i unless she is one of the stage 2 players and one of the first $t-1$ parties to speak has chosen not to reveal their share. In the latter case player i does not reveal her share either.

In addition to revealing her share, player i is required to send along the authentication information $\tau_i \in T$ that was provided to her by the dealer. In case that either the authentication fails, or that the player has refused to broadcast her message,[6] player i will be considered as having failed the authentication and chosen the special $\perp$ action.

After the two stages are completed, each player will locally use a reconstruction function $g_S \in g$ in order to try and compute the secret given the shares that have been revealed during the protocol's execution. By the properties of secret sharing, it follows that a party will be able to compute the secret (with certainty) at the end of the protocol if a set $S \subseteq N$ of at least $t-1$ *other* parties have revealed their shares, and otherwise she has no information about the secret (i.e. can compute it with probability only $1/|\Delta|$).

Protocol 4.3 (Reconstruction protocol)**.** *Let $(N, t, \Delta, \tilde{\Theta}, \tilde{\mu}, g)$ be an authenticated secret sharing scheme. We specify a corresponding reconstruction protocol $\Pi = (\Sigma, H, P, m^*, f^*)$ as follows:*

- $\Sigma = \left(\bigcup_{i \in N} \Theta_i \times T\right) \cup \{\perp\}$, *where* $\perp \notin \bigcup_{i \in N} \Theta_i \times T$,
- *the set H consists of all sequences $(a_1, \ldots, a_\ell) \in \Sigma^*$ such that $\ell \leq n$, $a_j \in \Sigma$ for all j.*
- *the set Z of terminal histories consists of all $h = (a_1, \ldots, a_n) \in H$ of length n.*
- *for every history $h \in H$ of length $i-1$, the next player function is defined by $P(h) = i$,*
- *for every non-terminal history $h = (a_1, \ldots, a_{i-1}) \in H \setminus Z$, and any $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j \neq i}) \in \tilde{\Theta}_i$, the next-message function of player $i = P(h)$ is defined as:*

$$m_i^*(\tilde{\theta}_i, h) = \begin{cases} (\theta_i, \tau_i) & \text{if } i < t, \\ \perp & \text{if } i \geq t \text{ and } \mathrm{Ver}_{\kappa_j}(a_j) = 0 \text{ for some } j < t, \\ (\theta_i, \tau_i) & \text{if } i \geq t \text{ and } \mathrm{Ver}_{\kappa_j}(a_j) = 1 \text{ for all } j < t. \end{cases} \tag{2}$$

- *for every terminal history $h = (a_1, \ldots, a_n) \in Z$, and every $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j \neq i}) \in \tilde{\Theta}_i$, the output function of player $i \in N$ is defined as*

$$f_i^*(\tilde{\theta}_i, h) = \begin{cases} g_R((a_j)_{j \in R_{-i}}, \theta_i) & \text{if } g_R((a_j)_{j \in R_{-i}}, \theta_i) \neq \lambda \text{ and } \mathrm{Ver}_{\kappa_j}(a_j) = 1 \text{ for all } j < t \\ \delta_0 & \text{otherwise} \end{cases}$$

where $R_{-i} = \{j \neq i : \mathrm{Ver}_{\kappa_j}(a_j) = 1\}$, $R = R_{-i} \cup \{i\}$, the g_T's are the reconstruction functions from the secret-sharing scheme, and δ_0 is an arbitrary element of Δ.

The protocol corresponds to the honest strategy in the secret sharing game. A crucial property of the honest strategy is that *even players that adhere to it do not necessarily reveal their share.* This will be the case when one of the first $t-1$ players aborts or deviates (by broadcasting $\perp$). On

[6] In an implementation, a player that fails to broadcast her value within some predetermined amount of time might be considered to have refused to broadcast.

the other hand, if none of the first $t-1$ players aborts or deviates, the honest strategy instructs to reveal, even if the history subsequent to the first $t-1$ rounds does contain an abort or a deviation.

Intuitively, allowing honest players to sometimes abort ensures that players who deviate from the protocol's prescribed strategy in early rounds (the first $t-1$) are penalized and unable to learn the secret themselves. On the other hand, the honest players will continue to report their share even if a deviation occurs in the t^{th} round; this ensures that the players in the first $t-1$ rounds will learn the secret in addition to the other players, and thus provides fairness.

The protocol requires that the players in each stage reveal their shares in sequential order. As we will argue in Section 5.1, the order in which the first $t-1$ players broadcast has no effect on the strategic properties of the protocol, and similarly for the last $n-t+1$ players. Thus, the protocol can effectively be implemented with two rounds of communication.

4.4 Rational Strategies for Corresponding Reconstruction Games

We consider reconstruction games that correspond to the reconstruction protocol Π with honest players. By our hypothesis, an induced reconstruction game is a game with public actions and private outputs $\Gamma = (N, H, P, A, L, \Delta, \tilde{\Theta} \times \Omega, \tilde{\mu} \times \zeta_m, u)$.

Our goal is to describe a set-valued rational strategy for the reconstruction game. The rational strategy instructs both honest and rational players to follow the strategy prescribed by Π, except that it does not specify how rational players should act in cases when the honest strategy may not be in their self interest. Specifically, we allow arbitrary action by the rational players when the first $t-1$ players have all revealed valid shares (whereas honest players must reveal in this case). The honest strategy (equivalent to the earlier 'protocol') is itself consistent with the rational set-valued strategy profile.

Definition 4.4 (Rational set-valued strategy)**.** *The* rational set-valued stategy $S_i = (M_i, F_i)$ *for player $i \in N$ in the game Γ is defined as follows:*

- *For every non-terminal history $h = (a_1, \ldots, a_{i-1}) \in H \setminus Z$ so that $i = P(h)$ and every $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j \neq i}) \in \tilde{\Theta}_i$, the public set-valued strategy of player i is defined as:*

$$M_i(\tilde{\theta}_i, \texttt{honest}, h) = \{m_i^*(\tilde{\theta}_i, h)\}$$

$$M_i(\tilde{\theta}_i, \texttt{rational}, h) = \begin{cases} \{(\theta_i, \tau_i)\} & \text{if } i < t, \\ \{\perp\} & \text{if } i \geq t \text{ and } \mathrm{Ver}_{\kappa_j}(a_j) = 0 \text{ for some } j < t, \\ \Sigma & \text{if } i \geq t \text{ and } \mathrm{Ver}_{\kappa_j}(a_j) = 1 \text{ for all } j < t. \end{cases}$$

- *For every terminal history $h = (a_1, \ldots, a_n) \in Z$ and every $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j \neq i}) \in \tilde{\Theta}_i$, the private set-valued strategy for player $i \in N$ is defined as:*

$$F_i(\tilde{\theta}_i, \texttt{honest}, h) = \{f_i^*(\tilde{\theta}_i, h)\}$$

$$F_i(\tilde{\theta}_i, \texttt{rational}, h) = \begin{cases} \Delta & \text{if } \mathrm{Ver}_{\kappa_j}(a_j) = 1 \text{ for all } j < t \\ \{\delta_0\} & \text{otherwise.} \end{cases}$$

where δ_0 is an arbitrary element of Δ.

Note that if all players follow any strategy s that is consistent with S and there exists an honest player $i \in N$ with $i \geq t$ then t out of the n players eventually reveal their share. Thus, if we can argue that: (1) it is likely to have an honest player $i \in N$ with $i \geq t$, and (2) rational players are incentivized to follow strategies consistent with S, then we will have obtained a fair solution to the secret sharing problem (since all players are likely to learn the secret following the protocol's execution). The way we envision an "implementation" of our protocol working is that players' machines are programmed with the honest strategy. Some (rational) players may decide to reprogram their own machines to behave differently on information sets where they have a chance of increasing their utility, and this may occur iteratively (as rational players react to what other rational players might do). Our set-valued equilibrium is meant to capture the idea that this process should never leave the set-valued strategy, and hence fairness will be maintained.

We start by proving that if all players follow a rational strategy $s \in S$ then everybody is likely to compute the secret correctly. The game-theoretic analysis is then provided in the next section.

Lemma 4.5 (Computing the secret)**.** *Suppose that players' types are chosen according to $\mu \times \zeta_m$, and that all players follow a strategy vector $s \in S$. Then the probability that all* honest *players compute the secret correctly in Γ is at least $1-(n-t+1)\gamma - p(n,m)$. If moreover s is a Nash equilibrium then the probability that* all *players compute the secret correctly in Γ is at least $1-(n-t+1)\gamma - p(n,m)$.*

Proof. Since the players follow a strategy $s \in S$, the first $t-1$ players will reveal their shares. So if there is a single honest player among the final $n-t+1$ players, the subsequent terminal history will contain at least t properly authenticated shares. The probability that this is not the case is $p(n,m)$. By the unforgeability of the authentication, the probability that any of the other messages contains a falsely authenticated share is at most $(n-t+1)\gamma$. Thus, $g_R((a_j)_{j\in R_{-i}}, \theta_i)$ equals the share with probability at least $1-(n-t+1)\gamma - p(n,m)$. In this case, all honest players are instructed to compute the secret, and will indeed succeed in doing so correctly.

Suppose now that the rational strategy $s \in S$ is a Nash equilibrium, and let $q = q(s,i)$ be the probability that player i does *not* compute the secret correctly with some probability (since $\mathrm{Ver}_{\kappa_j}(a_j) = 1$ for all $j < t$, we have that player i may choose to output any element in Δ in the private stage). Let s_i' be the honest strategy. As we saw before, with probability $1-(n-t+1)\gamma - p(n,m)$, the honest strategy s_i' enables player i to compute the secret correctly. By linearity of the utility function this means that, conditioned on players $j \neq i$ following s_{-i}, her expected utility in case she follows s_i' is larger than her expected utility in case she follows s_i by an additive amount of at least

$$\alpha_{ii}(q - (n-t+1)\gamma + p(n,m)).$$

Since $\alpha_{ii} > 0$, and s is a Nash Equilibrium, we conclude that $q \leq (n-t+1)\gamma + p(n,m)$. ■

5 Trembling-Hand Perfect Set Equilibrium

We prove that the rational set-valued strategy S is a trembling-hand perfect set equilibrium in the extensive game with public actions and private output Γ.

Lemma 5.1 (Trembling-hand perfect set equilibrium)**.** *The rational set-valued strategy vector $S = (M,F)$ from Definition 4.4 is a trembling-hand perfect set equilibrium in the game Γ.*

Proof: Our goal is to prove that S is a trembling-hand perfect set equilibrium for the game $\Gamma = (N, H, P, A, L, \Delta, \tilde{\Theta} \times \Omega, \tilde{\mu} \times \zeta, u)$.

By Lemma 2.7 it suffices to prove that for every $s \in S$, there exists a fully mixed sequence $(s^k)_{k=0}^{\infty}$ converging to s, so that for every player $i \in N$, every triplet $(\tilde{\theta}_i, \omega_i, h) \in \tilde{\Theta}_i \times \Omega_i \times H$ such that $P(h) = i$, and every strategy s_i'' for player i in the game $\Gamma(s^k, h)$ such that $s_i''(\tilde{\theta}_i, \omega_i, \epsilon) \notin S_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$, there is a strategy s_i' for player i such that $s_i'(\tilde{\theta}_i, \omega_i, \epsilon) \in S_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ and

$$u_i|_h(\mu|_h, s^k_{-i}|_h, s_i') \geq u_i|_h(\mu|_h, s^k_{-i}|_h, s_i'') \tag{3}$$

for all k. Note that it is sufficient to prove the statement for sufficiently large k, since then we can just start the sequence at a large enough value of k to handle all of the (finitely) many values of $i, \tilde{\theta}_i, \omega_i, h, s_i', s_i''$. Observe also that if $S_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ contains all possible actions available to player i (i.e., $M_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = A_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ for $h \in H \setminus Z$ and $F_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = L_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ for $h \in Z$), then what we need to prove holds trivially. This is because there does not exist $s_i''(\tilde{\theta}_i, \omega_i, \epsilon) \notin S_i(\tilde{\theta}_i, \omega_i, \epsilon)$.

Given $s \in S$ we define a sequence $(s^k)_{k=0}^{\infty}$ of strategy profiles in the following way. Fix any sequence (ε_k) of positive real numbers such that $\varepsilon_k \to 0$ as $k \to \infty$ (e.g. $\varepsilon_k = 1/k$). For any $i \in N$, any tuple $(\tilde{\theta}_i, \omega_i, h) \in \tilde{\Theta}_i \times \Omega_i \times H$, the strategy $s_i^k(\tilde{\theta}_i, \omega_i, h)$ equals $s_i(\tilde{\theta}_i, \omega_i, h)$ with probability $1 - \epsilon_k$ and is uniform over the action set $A_i(\tilde{\theta}_i, \omega_i, h)$ or $L_i(\tilde{\theta}_i, \omega_i, h)$ (i.e., which is a singleton in case $\omega_i = \texttt{honest}$ and is either Σ or Δ in case $\omega_i = \texttt{rational}$). It can be seen that $(s^k)_{k=0}^{\infty}$ converges to s as $k \to \infty$, and that for every $i \in N$ and k, the strategy s_i^k is fully mixed (because s_k hits every element in the action set with nonzero probability).

Let $i \in N$, let $h \in H$, such that $P(h) = i$, let $(\tilde{\theta}_i, \omega_i) \in \tilde{\Theta}_i \times \Omega_i$, and let s_i'' be any strategy for player i in game $\Gamma(s^k, h)$ such that $s_i''(\tilde{\theta}_i, \omega_i, \epsilon) \notin S_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$. We would like to show that there exists some s_i' such that $s_i'(\tilde{\theta}_i, \omega_i, \epsilon) \in S_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ for which Eq. (3) holds. Let $\tilde{\theta}_i = (\theta_i, \tau_i, \{\kappa_j\}_{j \neq i})$. We start by considering the case of a *nonterminal* history $h = (a_1, \ldots, a_{i-1})$. As noted above, there is nothing to prove when $M_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = A_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$. In particular, this always holds when $\omega_i = \texttt{honest}$ (because both sets are singletons) and it also holds when $\omega_i = \texttt{rational}$ and $M_i(\tilde{\theta}_i, \omega_i, h) = \Sigma$.

This leaves us with the cases in which $\omega_i = \texttt{rational}$ and $M_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = \{(\theta_i, \tau_i)\}$ (which occurs when $i < t$) and $M_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = \{\perp\}$ (which occurs when $i \geq t$ and $\mathrm{Ver}_{\kappa_j}(a_j) = 0$ for at least one $j < t$). In the latter case, all actions a', a'' give player i the same utility, because the reconstruction strategies (and hence utilities under s^k) do not depend on the action of player i or any subsequent player. Specifically, if h is any terminal history in which one of the first $t-1$ players do not reveal valid shares then $F_\ell|_h(\tilde{\theta}_\ell, \texttt{honest}, \epsilon) = F_\ell|_h(\tilde{\theta}_\ell, \texttt{rational}, \epsilon) = \{\delta_0\}$, so $s_\ell(\tilde{\theta}_\ell, \omega_\ell, h) = \delta_0$, and $s_\ell^k(\tilde{\theta}_\ell, \omega_\ell, h)$ will equal δ_0 with probability $1 - \epsilon_k$ and will be uniform on either $\{\delta_0\}$ or Δ with probability ϵ_k.

We are left with the case in which $\omega_i = \texttt{rational}$ and $i < t$ (which is at the heart of our argument). We need to show that outputting (θ_i, τ_i) on $(\tilde{\theta}_i, \omega_i, \epsilon)$ in subgame $\Gamma(s^k, h)$ is a better response to s^k_{-i} than any strategy s_i'' that does not do this.

Claim 5.2. *Suppose that $\omega_i = \texttt{rational}$ and $i < t$. Let $s_i' = (m_i', f_i')$ be any strategy in game $\Gamma(s^k, h)$ such that m_i' outputs (θ_i, τ_i) on input $(\tilde{\theta}_i, \omega_i, \epsilon)$ and s_i' agrees with $s_i^k|_h$ everywhere else, and let s_i'' be any strategy in game $\Gamma(s^k, h)$ such that $s_i''(\tilde{\theta}_i, \omega_i, \epsilon) \neq (\theta_i, \tau_i)$. Then $u_i|_h(\bar{\mu}|_h, s^k_{-i}|_h, s_i') \geq u_i|_h(\bar{\mu}|_h, s^k_{-i}|_h, s_i'')$.*

Proof of Claim: First suppose that $\mathrm{Ver}_{\kappa_j}(a_j) = 0$ for at least one $j < i$. Then, as above, the reconstruction strategies (and hence utilities) do not depend on the action of player i or any subsequent player, and thus the claim holds.

So now suppose that $\mathrm{Ver}_{\kappa_j}(a_j) = 1$ for all $j < i$. There are two ways that a previous player j could have produced an authenticated share a_j under strategy s^k. It could have

sent $a_j = (\theta_j, \tau_j)$ using the "un-trembled" portion of the strategy, and it could have sent a_j using the "trembled" portion of the strategy that is uniformly random on Σ. Note that the trembled portion of the strategy is used with probability at most ε_k (indeed at most $(1 - m/n) \cdot \varepsilon_k$, since only rational players have the possibility of trembling). Moreover, by our definition of admissible secret sharing, these trembled and untrembled strategies for the first $i-1$ players are identically distributed once we condition on the shares being valid wrt $\{\kappa_j\}_{j \neq i}$ (by Definition 3.2, Item 1). This implies that conditioned on history $(\tilde{\theta}_i, h)$, the probability that all of the previous actions a_j are equal to the un-trembled shares θ_j is at least $1 - \varepsilon_k \cdot (i-1)$.

Now, suppose that player i reveals her authenticated share (θ_i, τ_i). Then, player i can compute the secret correctly provided that none of the previous actions are trembles, none of the future actions are trembles, and at least one of the last $n - t$ players is honest. We have already argued that, conditioned on $(\tilde{\theta}_i, h)$, the probability that one of the previous actions was a tremble is at most $\varepsilon_k \cdot (i-1)$. The probability of a future tremble is at most $\varepsilon_k \cdot (n-i)$. The probability that there are no honest players among the last $n - t$ is precisely $p(n, m)$. In all, player i computes the secret correctly with probability at least $1 - p(n,m) - \varepsilon_k \cdot (n-1)$. Thus her expected utility conditioned on $(\tilde{\theta}_j, h)$ is at least $(1 - p(n,m) - \varepsilon_k \cdot (n-1)) \cdot \alpha_{ii} + \sum_{j \neq i} \alpha_{ij}$. Here we use the fact that strategy f_i' maximizes player i's expected utilities amongst all strategies in F_i, including any strategy that attempts to compute the secret.

Suppose instead that player i sends a different message $a'' \neq (\theta_i, \tau_i)$. By the unforgeability of the authenticated secret-sharing scheme, the probability that a'' passes verification is at most γ. If it is invalid, then the public actions of all future players will be independent of their shares (they will either be $\perp$ or uniformly random trembles), and thus player i's probability of computing the secret correctly will remain $1/|\Delta|$ (as it is conditioned on history h). Other players will have private outputs that are either δ_0 or a random tremble of this, and hence will also compute the secret correctly with probability $1/|\Delta|$. Thus player i's expected utility conditioned on $(\tilde{\theta}_j, h)$ is $(1/|\Delta|) \cdot \sum_j \alpha_{ij}$.

Player i strictly prefers to send $a' = (\theta_i, \tau_i)$ provided that

$$(1 - p(n,m) - \varepsilon_k \cdot (n-1)) \cdot \alpha_{ii} + \sum_{j \neq i} \alpha_{ij} > (1/|\Delta|) \cdot \sum_j \alpha_{ij}.$$

This holds for sufficiently large k by our hypothesis (1) on the players' preference for learning the secret. ■

We now turn to consider the case of a terminal history $h = (a_1, \ldots, a_n)$. As noted above, there is nothing to prove in the case where $F_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = L_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$. This always holds when $\omega_i = \texttt{rational}$ and $F_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = \Delta$, and when $\omega_i = \texttt{honest}$ (since in this case $F_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ and $L_i|_h(\tilde{\theta}_i, \omega_i, \epsilon)$ are both singletons). So we are left with the case in which $\omega_i = \texttt{rational}$ and $F_i|_h(\tilde{\theta}_i, \omega_i, \epsilon) = \{\delta_0\}$ (which occurs if $\mathrm{Ver}_{\kappa_j}(a_j) = 0$ for some $j < t$).

Claim 5.3. *Suppose that $\omega_i = \texttt{rational}$ and $\mathrm{Ver}_{\kappa_j}(a_j) = 0$ for some $j < t$. Let s_i' be a strategy in game $\Gamma(s^k, h)$ that outputs $\delta_0 \in \Delta$ on $(\tilde{\theta}_i, \omega_i, \epsilon)$ and agrees with $s_i^k|_h$ everywhere else, and let s_i'' be any other strategy. Then $u_i|_h(\bar{\mu}|_h, s_{-i}^k|_h, s_i') \geq u_i|_h(\bar{\mu}|_h, s_{-i}^k|_h, s_i'')$.*

Proof of Claim: Since $\mathrm{Ver}_{\kappa_j}(a_j) = 0$ for some $j < t$, we have that for all $\ell > t$, the strategy s_ℓ instructs player $\ell > t$ to send $\perp$ (whether $\omega_\ell = \texttt{honest}$ or $\omega_\ell = \texttt{rational}$).

So the only way the history h can include messages other than $\perp$ for players $\ell \geq t$ is via the random trembles of s^k_ℓ. Since the trembles are independent of the shares θ_ℓ, they do not reveal any information about them. This remains true even if the trembles happen to pass the authentication. In addition, by the secrecy property of the secret sharing scheme (last item in Definition 3.1), the first $t-1$ messages do not reveal any information about the secret to any of the players. Thus, conditioned on $(\tilde{\theta}_i, h)$ player i has no information about the secret. Since s^k_{-i} instructs all other players to output δ_0 or a uniformly distributed element of Δ, and given that the secret δ is uniformly distributed in Δ (conditioned on $(\tilde{\theta}_i, h)$), we have that any action $\delta'' \in \Delta$ will yield player i the same expected utility as δ_0, namely $(1/|\Delta|) \cdot \sum_j a_{ij}$. ■

■

Remark 5.1. *As discussed earlier, the honest-player distribution ζ_m in Theorem 4.2 models the situation where players have no a priori information about which players are honest or rational, except the probability (m/n) with which an individual player is honest. An opposite extreme is the case where the set of honest players is not random, but is instead an arbitrary fixed set of m players, known to all. Our protocol can be modified to handle this case by first randomly permuting the order in which players speak; this guarantees that there will be an honest player among $t, \ldots, n$ with high probability. Now the rational strategy would have the first $t-1$ players determine whether the permutation is 'good', reveal their share if it is, and send $\perp$ otherwise. This analysis can be found in the preliminary version of our paper [38], where it is also shown to be coalition-proof in the sense of Bernheim et al. [4]. By combining these ideas, it may be possible to handle arbitrary distributions on the set of honest players (provided there are at least roughly m honest players with high probability).*

5.1 Modeling Rushing

So far the reconstruction protocol has specified a particular sequence with which the first $t-1$ players each reveal their shares, and similarly for the subsequent sequencing of the remaining $n-(t-1)$ players. But the protocol satisfies a robustness property: the strategy prescribed to both rational players and honest players in both stage 1 (for the first $t-1$ players) and in stage 2 (for the remaining $n-t+1$ players) is invariant to the action selected by other players within the same stage. This makes the equilibrium robust, also, to a simpified two round protocol in which $t-1$ players are selected to send a message in the first round and the remaining $n-(t-1)$ players in the second round. The communication channel within each round need not provide for simultaneous broadcast, and can be just a synchronous broadcast channel that allows for "rushing" where a player can wait and see the message sent by another player before choosing its own message. Not sending a message during a round is interpreted as $\perp$.

Formal analysis of this rushing game could be provided as a partially-specified game in the spirit of Kalai [25], wherein the same equilibrium survives many extensive variations of a two-round, simultaneous-move Bayesian game-form for the protocol. This two-round simultaneous-move secret-sharing game would ask each of a random set $t-1$ of players to broadcast a message on a *synchronous* communication channel. Each of the remaining $n-(t-1)$ players would then be asked to broadcast a message on the synchronous communication channel. Finally, all players would select a private action just as in our n-round protocol. The strategy described earlier remains an equilibrium in this two-round game, and also for any sequencing of play within the two rounds. Kalai develops a general theory of structural robustness in which equilibrium are robust to variations in

details about a game, for the setting of large semi-anonymous games that satisfy a payoff continuity property. We achieve similar robustness properties in our secret-sharing game.

6 Acknowledgements

We thank Drew Fudenberg, Jonathan Katz, Silvio Micali, Peter Bro Miltersen, Moni Naor, Adam O'Neill and the anonymous reviewers for helpful discussions.

References

[1] *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing, 2-4 May 1988, Chicago, Illinois, USA*. ACM, 1988.

[2] I. Abraham, D. Dolev, R. Gonen, and J. Y. Halpern. Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation. In E. Ruppert and D. Malkhi, editors, *PODC*, pages 53–62. ACM, 2006.

[3] A. S. Aiyer, L. Alvisi, A. Clement, M. Dahlin, J.-P. Martin, and C. Porth. Bar fault tolerance for cooperative services. In A. Herbert and K. P. Birman, editors, *SOSP*, pages 45–58. ACM, 2005.

[4] B. P. B. Douglas Bernheim and M. D. Whinston. Coalition-proof nash equilibria i. concepts. *Journal of Economic Theory*, 42(1):1–12, 1987.

[5] M. Babaioff, R. Lavi, and E. Pavlov. Mechanism design for single-value domains. In *Proc. Nat. Conf. on Artificial Intelligence, AAAI05*, 2005.

[6] K. Basu and J. W. Weibull. Strategy subsets closed under rational behavior. *Economics Letters*, 36:141–146, 1991.

[7] M. Ben-Or, S. Goldwasser, and A. Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In *STOC* [1], pages 1–10.

[8] G. Blakely. Safeguarding cryptographic keys. In *AFIPS*, volume 48, page 313, 1979.

[9] D. Boneh and M. Naor. Timed commitments. In *Proc. CRYPTO 2000*, pages 236–254, 2000.

[10] R. Canetti, editor. *Theory of Cryptography, Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19-21, 2008*, volume 4948 of *Lecture Notes in Computer Science*. Springer, 2008.

[11] D. Chaum, C. Crépeau, and I. Damgård. Multiparty unconditionally secure protocols (extended abstract). In *STOC* [1], pages 11–19.

[12] B. Chor, S. Goldwasser, S. Micali, and B. Awerbuch. Verifiable secret sharing and achieving simultaneity in the presence of faults (extended abstract). In *FOCS*, pages 383–395. IEEE, 1985.

[13] R. Cleve. Limits on the security of coin flips when half the processors are faulty (extended abstract). In *STOC*, pages 364–369. ACM, 1986.

[14] G. B. Davis and T. W. Sandholm. Algorithms for Rationalizability and CURB Sets. In *AAAI'06*, 2006.

[15] S. Even, O. Goldreich, and A. Lempel. A randomized protocol for signing contracts. *Commun. ACM*, 28(6):637–647, 1985.

[16] J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker. A BGP-based mechanism for lowest-cost routing. In *Proceedings of the 2002 ACM Symposium on Principles of Distributed Computing*, pages 173–182, 2002.

[17] J. Feigenbaum, C. H. Papadimitriou, and S. Shenker. Sharing the cost of multicast transmissions. *Journal of Computer and System Sciences*, 63:21–41, 2001.

[18] J. Feigenbaum and S. Shenker. Distributed Algorithmic Mechanism Design: Recent Results and Future Directions. In *Proceedings of the 6th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 1–13, 2002.

[19] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

[20] J. A. Garay and M. Jakobsson. Timed release of standard digital signatures. In *Proc. Financial Cryptography 2002*, pages 168–182, 2002.

[21] O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game or a completeness theorem for protocols with honest majority. In *STOC*, pages 218–229. ACM, 1987.

[22] S. D. Gordon and J. Katz. Rational secret sharing, revisited. In R. D. Prisco and M. Yung, editors, *SCN*, volume 4116 of *Lecture Notes in Computer Science*, pages 229–241. Springer, 2006.

[23] J. Y. Halpern and V. Teague. Rational secret sharing and multiparty computation: extended abstract. In L. Babai, editor, *STOC*, pages 623–632. ACM, 2004.

[24] S. Izmalkov, S. Micali, and M. Lepinski. Rational secure computation and ideal mechanism design. In *FOCS*, pages 585–595. IEEE Computer Society, 2005.

[25] E. Kalai. Large robust games. *Econometrica*, 72(6):1631–1665, November 2004.

[26] J. Katz. Bridging game theory and cryptography: Recent results and future directions. In Canetti [10], pages 251–272.

[27] G. Kol and M. Naor. Cryptography and game theory: Designing protocols for exchanging information. In Canetti [10], pages 320–339.

[28] G. Kol and M. Naor. Games for exchanging information. In R. E. Ladner and C. Dwork, editors, *STOC*, pages 423–432. ACM, 2008.

[29] D. M. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, 50:863–894, 1982.

[30] R. Lavi and N. Nisan. Online ascending auctions for gradually expiring goods. In *SODA'05*, 2005.

[31] D. Lehmann, L. I. O'Callaghan, and Y. Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM*, 49(5):577–602, September 2002.

[32] M. Lepinski, S. Micali, C. Peikert, and A. Shelat. Completely fair sfe and coalition-safe cheap talk. In S. Chaudhuri and S. Kutten, editors, *PODC*, pages 1–10. ACM, 2004.

[33] M. Lepinski, S. Micali, and A. Shelat. Collusion-free protocols. In H. N. Gabow and R. Fagin, editors, *STOC*, pages 543–552. ACM, 2005.

[34] A. Lysyanskaya and N. Triandopoulos. Rationality and adversarial behavior in multi-party computation. In C. Dwork, editor, *CRYPTO*, volume 4117 of *Lecture Notes in Computer Science*, pages 180–197. Springer, 2006.

[35] R. McGrew, R. Porter, and Y. Shoham. Towards a general theory of non-cooperative computation. In J. Y. Halpern and M. Tennenholtz, editors, *TARK*, pages 59–71. ACM, 2003.

[36] N. Nisan and A. Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35:166–196, 2001.

[37] A. O'Neill and A. Sangwan. Honesty, rationality, and malice in secret sharing and mpc: Robust protocols for real-world populations. In *Manuscript*, 2008.

[38] S. J. Ong, D. Parkes, A. Rosen, and S. Vadhan. Fairness with an honest minority and a rational majority. Available from `http://eecs.harvard.edu/~salil/Fairness-abs.html`, April 2007.

[39] S. J. Ong, D. Parkes, A. Rosen, and S. Vadhan. Fairness with an honest minority and a rational majority. Cryptology ePrint Archive, Report 2008/097, March 2008. `http://eprint.iacr.org/`.

[40] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[41] D. C. Parkes and J. Shneidman. Distributed implementations of Vickrey-Clarke-Groves mechanisms. In *Proc. 3rd Int. Joint Conf. on Autonomous Agents and Multi Agent Systems*, pages 261–268, 2004.

[42] A. Petcu, B. Faltings, and D. Parkes. M-dpop: Faithful distributed implementation of efficient social choice problems. In *AAMAS'06 - Autonomous Agents and Multiagent Systems*, pages 1397–1404, Hakodate, Japan, May 2006.

[43] B. Pinkas. Fair secure two-party computation. In *Proc. EUROCRYPT 2003*, pages 87–105, 2003.

[44] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority (extended abstract). In *STOC*, pages 73–85. ACM, 1989.

[45] R. Selten. A reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.

[46] A. Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.

[47] J. Shneidman and D. C. Parkes. Specification faithfulness in networks with rational nodes. In *Proc. 23rd ACM Symp. on Principles of Distributed Computing (PODC'04)*, St. John's, Canada, 2004.

[48] Y. Shoham and M. Tennenholtz. Non-cooperative computation: Boolean functions with correctness and exclusivity. *Theor. Comput. Sci.*, 343(1-2):97–113, 2005.

[49] M. N. Wegman and L. Carter. New hash functions and their use in authentication and set equality. *J. Comput. Syst. Sci.*, 22(3):265–279, 1981.

[50] A. C.-C. Yao. How to generate and exchange secrets (extended abstract). In *FOCS*, pages 162–167. IEEE, 1986.